

ESTRA: UM *CORPUS* PARA O ESTUDO DO ESTILO DA TRADUÇÃO

Célia M. Magalhães*
Universidade Federal de Minas Gerais

Resumo: Este artigo apresenta a evolução e as contribuições da pesquisa em estudos da tradução orientados para *corpora* no Brasil. Faz-se uma revisão dos trabalhos iniciais desenvolvidos no Laboratório Experimental de Tradução (LETRA), mostrando que a maioria deles adotava uma abordagem de linguística contrastiva da tradução e que as pesquisas, gradualmente, foram evoluindo para uma preocupação com a estilística tradutória e o estilo do tradutor literário. Também se relata a compilação de um *corpus* para o estudo do estilo da tradução, o ESTRA, projetado exclusivamente para esse fim. Mostra-se como a pesquisa de *corpora* do ESTRA promove a interdisciplinaridade nos estudos da tradução e introduz a triangulação de resultados de análises realizadas com procedimentos metodológicos das diferentes abordagens usadas para estudar o estilo. Descrevem-se procedimentos metodológicos novos, em especial a etiquetagem do *corpus* para algumas das categorias de estilo. Termina-se o artigo com uma visão crítica sobre o que tem sido feito até o presente, apresentando perspectivas futuras de pesquisa em estilística tradutória no LETRA.

Palavras-chave: Estilo da tradução. Estilo do tradutor. Pesquisa orientada por *corpora*.

Célia M. Magalhães. Possui graduação em Letras/Licenciatura em Inglês e mestrado em Letras/Linguística pela Universidade Federal de Minas Gerais, Brasil. Doutorou-se em Literatura Comparada pela Universidade Federal de Minas Gerais, em 1997. Atualmente é professora Titular da Faculdade de Letras da Universidade Federal de Minas Gerais. E-mail: cmagalhaes@ufmg.br.



ESTRA: A CORPUS FOR THE STUDY OF THE STYLE OF TRANSLATION

Abstract: This article aims at presenting the developments and contributions of corpus-oriented research in Brazil. It first reviews initial work carried out at the Laboratory of Experimentation in Translation (LETRA), at the Federal University of Minas Gerais (UFMG). It then shows most of this research has taken a contrastive linguistic approach to translation and gradually evolved to a concern with translational stylistics and the literary translator's style. The article also reports on the compilation of a corpus for the study of style of translation (ESTRA), designed exclusively for that purpose. It shows how the ongoing corpus work developed with ESTRA enhances interdisciplinarity in TS and introduces triangulation of data obtained through methodological procedures from the different research strands used to approach style. It describes the new methodologies used, especially for tagging the corpus for style categories. The article ends with a critical view on what has been done so far and tries to envisage future perspectives for translational stylistics at LETRA.

Keywords: Style of translation. Translator's style. Corpus-oriented research.

Introdução

Partindo de trabalhos dos estudos da tradução publicados no contexto internacional que utilizam *corpus* como parte de sua metodologia, em especial trabalhos dos estudos da tradução baseados em *corpus* (ETBC), este artigo tem como objetivo mostrar alguns desdobramentos da pesquisa em tradução orientada para *corpora* no contexto nacional. Dá-se enfoque a um relato dos trabalhos iniciais desenvolvidos no âmbito do Laboratório de Experimental de Tradução (LETRA), da Universidade Federal de Minas Gerais (UFMG).

Os trabalhos aqui destacados têm sua origem em uma primeira fase de uso de *corpora* paralelos com uma abordagem linguística contrastiva; passam por uma segunda fase em que se exploram os *corpora* a partir de abordagens da semiótica social, introduzindo elementos do contexto sociocultural mais amplo das traduções; e, finalmente, chegam a uma terceira fase, em que se afiliam aos

estudos de estilo do texto traduzido (TT) e, em menor dimensão, do tradutor, com a criação e utilização do Corpus de Estilo da Tradução (ESTRA). O artigo objetiva, ainda, mostrar as contribuições originais referentes à própria compilação do ESTRA e às metodologias usadas no âmbito desse *corpus*, além de perspectivas futuras das pesquisas sobre estilo no contexto nacional.

Este artigo está organizado em cinco seções, incluindo esta Introdução e as Considerações Finais. A segunda seção faz uma retrospectiva dos primeiros trabalhos sobre tradução com orientação para *corpus*, desenvolve os conceitos de estilo usados no âmbito dos estudos da tradução e relata o *status* atual dos trabalhos realizados com o ESTRA. A terceira seção descreve os procedimentos de compilação e preparação dos textos do ESTRA para análises (semi) automáticas. Nas Considerações Finais, faz-se uma revisão crítica dos resultados obtidos até o momento e apontam-se as perspectivas futuras para a pesquisa de estilo.

1. Revisão da literatura

1.1 Retrospectiva: os estudos da tradução e o uso de corpora

Baker (1993, 1995) são os textos clássicos de introdução dos ETBC, linha de pesquisa criada a partir da necessidade de descrever o TT como evento comunicativo genuíno, usando grande quantidade de textos eletrônicos para análise (semi)automática. A descrição dos TTs inicialmente se voltou à identificação/corroboração de universais da tradução, os quais, posteriormente, passaram a ser chamados de características ou, simplesmente, traços do TT inerentes ao processo de tradução. Vários trabalhos avançaram na direção da explicação de traços recorrentes dos textos traduzidos, como a simplificação, a explicitação e a normalização, todos definidos em Baker (1993, 1995), que apresentam, ainda, as primeiras

definições dos diferentes tipos de *corpus*, mais tarde refinadas, dentre outros autores, por Kenny (2001) e Zanettin (2012).

Faz parte do desenvolvimento de uma linha de pesquisa orientada para *corpora* no Brasil a compilação do CORDIAL – Corpus Discursivo para Análises Linguísticas e Literárias –, iniciada no LETRA em torno de 1999. O CORDIAL constituía-se majoritariamente como um *corpus* paralelo com a predominância do par linguístico inglês/português brasileiro (PB) ou europeu (PE), mas também com exemplares dos pares francês/PB, espanhol/PB, PB/inglês e PB/italiano. Pode-se dizer que a maioria das pesquisas realizadas com textos desse *corpus* (cf. MAURI, 2003; ASSIS, 2004, entre outros) adotava uma abordagem linguística da tradução, no sentido que é dado a esse rótulo em revisões críticas sobre o tema em coletâneas de vários tipos de estudos da tradução (cf. MAL-MKJÆR, 2011, entre outros). Os textos eram etiquetados com categorias da linguística sistêmico-funcional (LSF) para fins de análises das equivalências e mudanças nas traduções para o português. Os resultados dos trabalhos levaram a uma reflexão crítica de que a inexistência de uma descrição do PB ou do PE que fizesse uso dos pressupostos teóricos e metodológicos da LSF constituía uma restrição à análise contrastiva envolvendo textos nessas duas línguas, cujos usos são diferentes daqueles do inglês (HALLIDAY, 1994).

Como exceções a esse tipo de abordagem podem-se citar Magalhães (2005) e Bueno (2005). Magalhães (2005), ainda abordando o *corpus* com categorias de coesão lexical predominantemente calcadas na LSF, é uma investigação que se baseia igualmente em teóricos da tradução como Hatim e Mason (1992), que abordam a questão do estilo como escolha deliberada para significar e a questão do desvio como destaque (*foregrounding*) (ver mais à frente definições desses termos). Bueno (2005) é uma aplicação da metodologia usada em Kenny (2001) para investigar o léxico, a criatividade e potenciais normalizações na tradução de *Macunaíma* para o inglês.

A partir de 2009, observa-se uma significativa mudança de abordagem, com destaque para Assis (2009), Magalhães e Assis

(2009), Mauri (2009) e Bueno (2011). Os dois primeiros abordam *corpora* paralelos inglês/PB a partir da teoria de representação de atores sociais, uma teoria semiótica social crítica. Ambos focalizam aspectos diferentes de duas traduções do mesmo texto-fonte (TF), *Heart of darkness* (HOD), de Joseph Conrad, e os resultados apresentam questões latentes de estilo do texto, no sentido de escolhas deliberadas com prováveis construções diferentes de significados nos TTs em PB. Mauri (2009) investiga duas traduções italianas de Clarice Lispector produzidas por Adelina Aletti, *Legami familiari* e *L'ora della stella*, com enfoque no ponto de vista narrativo e nos efeitos de mudanças da apresentação do pensamento das personagens nos TFs para o fluxo de pensamento nos TTs. Bueno (2011) analisa um *corpus* de traduções para o PB e para o espanhol de *Chapeuzinho Vermelho/Caperucita Roja*, sem a preocupação de compará-las extensivamente com o TF. As análises de Bueno (2011) envolvem perspectivas distintas entre si, com teorias da semiótica social, sobre a representação (visual) de atores e ações sociais, e abrangem as capas das traduções como parte do contexto de sua produção. O interesse é voltado para a escolha deliberada de signos verbais e visuais que constroem significados diferentes no contexto cultural das traduções. Os trabalhos dessa segunda fase de pesquisa foram a motivação para a compilação de um novo *corpus* para o estudo do estilo da tradução, que veio a ser denominado ESTRA, Corpus de Estilo da Tradução.

Chegou-se, então, a algumas decisões: em primeiro lugar, dedicar-se à linha de pesquisa do estilo da tradução, ainda incipiente no país, restrita, até onde se sabe, a Camargo (2004), entre outros, e Aguiar (2010). Na linha de Malmkjær (1998, p. 539), que sugere como um avanço para os estudos de *corpus* de tradução a criação de *corpora* paralelos de pequeno porte, com várias traduções de um mesmo TF, para suplementar o grande número de estudos quantitativos com análise textual interpretativa, “se quisermos chegar pelo menos perto de ser capaz de dizer alguma coisa sobre como as ideias/conceitos/informações são realizadas em línguas diferentes”, decidiu-se compilar um novo *corpus*. Em segundo lu-

gar, supôs-se que, além de poder dizer como “ideias/conceitos/informações são realizadas em línguas diferentes”, a comparação entre TTs para uma mesma língua teria algo a dizer sobre as escolhas deliberadas para significar em contextos culturais e temporais diferentes daqueles do TF, concordando com Munday (2008) que o TF seria a variável imutável em relação aos TTs.

1.2 Estilo da tradução

Nida (1964) e Vinay e Darbelnet ([1958]1995) podem ser considerados trabalhos que tangenciam a questão de estilo, o segundo abordando traços de estilo como procedimentos de tradução. Hatim e Mason (1992, p. 10) adotam o conceito de estilo como: “resultado das escolhas motivadas feitas pelos produtores dos textos”. Abordam paralelismo, coesão e outros traços textuais como traços do TF potencialmente motivados e a serem levados em conta na tradução. Desvio, proeminência e relevância literária são conceitos da estilística que teóricos como Leech e Short ([1981]2007) elaboram a partir da noção de destaque (*foregrounding*) da Escola Linguística de Praga e de uma orientação hallidayana de análise do estilo. Leech e Short ([1981]2007) apresentam as seguintes definições: desvio é uma noção estatística que aponta a diferença entre a frequência usual de uma palavra e sua frequência em um texto ou *corpus* [de referência]; proeminência é a noção psicológica relacionada ao desvio, referindo-se ao modo como um traço linguístico se destaca para o analista ou leitor; e relevância literária é uma característica do desvio quando artisticamente motivado. O destaque (*foregrounding*) – conceito também psicológico, pois relacionado ao leitor/analista do texto – refere-se à percepção pelo leitor de um desvio no texto. Na acepção hallidayana, destaque e proeminência deixam de ter caráter psicológico e podem referir-se também a padrões consistentes da linguagem relacionados com o significado do texto como um todo, ou “proeminência motivada” (Halliday, 1971: 339).

Um dos primeiros trabalhos a propor uma metodologia específica para o estudo do estilo do tradutor é Baker (2000). Sua definição de estilo é a seguinte:

Entendo estilo como um tipo de impressão digital do polegar que é expressa em uma gama de traços linguísticos – bem como não linguísticos. [...] é uma questão de padrão: envolve a descrição de padrões preferidos ou recorrentes de comportamento linguístico, em lugar de instâncias individuais ou únicas de intervenção. (BAKER, 2000, p. 245)

Essa definição tem afiliações com a linguística forense e tem como interesse o estilo do tradutor – estilo esse cujo estudo é baseado em *corpora* comparáveis de textos traduzidos e não traduzidos.

Saldanha (2011) faz uma revisão crítica abrangente da pesquisa sobre estilo no âmbito dos estudos da tradução, distinguindo as que se ocupam do estilo do tradutor daquelas que têm como foco o estilo da tradução e que usam *corpora* paralelos com esse objetivo. Como o interesse neste artigo é pelos *corpora* paralelos, volta-se para os trabalhos que usam esses *corpora* e têm o estilo do texto traduzido como objeto de estudo.

Em Scott (1998), pode-se observar o interesse no estilo do texto traduzido e nos possíveis efeitos em sua recepção pelos leitores da cultura receptora. Em sua análise de um *corpus* com o TF *A hora da estrela* (Clarice Lispector) e o TT *The hour of the star* (Giovanni Potiero), Scott (1988) observa, por um lado, padrões de escolha de léxico de negação (*nada, não*, entre outros) no TF que constroem um dos temas da obra, a impotência da personagem *Macabéa*, e, por outro, escolhas do tradutor que não reproduzem esses padrões, normalizando algumas das escolhas artisticamente motivadas.

Kenny (2001), por sua vez, investiga os *hapax legomena* (palavras que ocorrem uma vez em um *corpus*) e seu potencial como palavras criativas, examinando textos literários experimentais traduzidos do alemão para o inglês. Tais escolhas criativas, indícios

do estilo dos textos do *corpus*, são até certo ponto normalizadas, mas são evidenciadas também soluções criativas nos TTs.

A definição de uma metodologia intitulada “estilística tradutória” é dada por Malmkjær (2003, 2004). Malmkjær (2004, p. 14) discute a multiplicidade de sentidos do termo “estilo” e cita Malmkjær e Carter (2002, p. 510), em que a noção de estilo é equacionada com a noção de registro ou tipo textual: “Estilo” pode ser definido como uma regularidade consistente e estatisticamente significativa de ocorrência no texto de certos itens e estruturas, entre aqueles oferecidos pela língua como um todo [...].

Malmkjær (2004, p. 14-15) discute, ainda, dois tipos de abordagem do estilo de textos, orientada para o escritor e orientada para o leitor. O interesse dela volta-se para a primeira abordagem, pois, no caso da tradução, o estudo do estilo, se orientado para o tradutor como escritor, deve observar que o tradutor é um mediador. Como um texto mediado, o TT é afetado por parâmetros que a autora define como: (a) a interpretação do tradutor/leitor; (b) o novo propósito para escrever o TT; (c) a diferença de propósitos do escritor do TF e do escritor da tradução; e (d) os públicos-alvo diferentes do TF e do TT. Para Malmkjær (2004, p. 16), o TT, portanto, só pode ser abordado de forma produtiva se for usada o que denomina de “estilística tradutória”, definida como uma metodologia que

... leva em consideração a relação entre o texto traduzido e seu texto-fonte [...] [para saber] [...] não apenas *como* o texto significa o que significa, mas também *por que* o escritor pode ter escolhido dar forma ao texto de um modo particular para fazê-lo significar do modo como ele significa [...]. (MALMKJÆR, 2004, p. 14, referindo-se a LEECH e SHORT, [1981]2007, p. 14)

Em busca de uma explicação no contexto histórico-cultural mais amplo, Malmkjær (2003, 2004) apresenta exercícios de estilística tradutória por meio de inúmeros exemplos de um *corpus* de TTs

em inglês produzidos por Henry William Dulcken a partir da obra mais ampla do escritor dinamarquês Hans Christian Andersen. Esse tipo de *corpus*, com TTs de um mesmo tradutor da obra de um autor, é considerado por Malmkjær como ideal para a análise da estilística tradutória, mas a autora também sugere o uso de TTs produzidos por tradutores diferentes a partir de um mesmo TF. A explicação de certas escolhas dos TTs de Dulcken usa, ainda, os subsídios dos parâmetros que afetam a tradução como texto mediado, indicados acima. Reconhecendo que não encontrou evidências para as explicações sugeridas, talvez pela falta de registros de reflexões do tradutor, Malmkjær (2004) aponta que as mudanças de estilo das traduções podem ter sido motivadas pelo desejo de Dulcken de atender a um público vitoriano do século XIX, especialmente quanto a ideias religiosas muito distintas construídas nos textos de Andersen.

Munday (2007, 2008) acrescenta a questão da ideologia, seguindo uma orientação dos trabalhos da análise crítica do discurso (ACD), para explicar, com base na voz do tradutor, parte das escolhas de estilo. Munday (2008), fazendo interface entre estudos descritivos da tradução, narratologia, metodologia de *corpus* e categorias da LSF tradicionalmente usadas em algumas das abordagens da ACD, propõe um quadro metodológico para análise, em dois níveis, do estilo da tradução e da voz, ou presença discursiva, do tradutor: (i) o contexto de situação/registo, com as funções ideacional, interpessoal e textual; e (ii) as escolhas léxico-gramaticais dos TTs, algumas delas, consoante Munday (2008), mais conscientes ou deliberadas e apontando na direção de ideologias na realização de significados motivados nos TTs. Essas escolhas realizam os seguintes planos do ponto de vista narrativo segundo Simpson (2000): psicológico, em interface com a função ideacional e realizado especialmente por escolhas da transitividade; ideológico, em interface com a função interpessoal e realizado especialmente por escolhas da modalidade; e espaço-temporal, em interface com a função textual e realizado por escolhas da estrutura temática e da dêixis. Munday (2008) observa que os planos se sobrepõem, assim

como os tipos de escolhas que os realizam. Acrescenta, ainda, o plano fraseológico e argumenta a favor do estudo da “primazia”, que poderia ser base para o estudo da presença discursiva do tradutor em escolhas linguísticas mais inconscientes nesse plano. Nesse ponto, a abordagem de estilo de Munday volta-se também para o estilo do tradutor, aproximando-se daquela de Baker (2000). Por fim, Saldanha (2011), já apontada anteriormente, apresenta uma tentativa de conciliar características do hábito linguístico do tradutor (estilo como atributo pessoal) com escolhas retóricas deliberadas no texto (estilo como atributo textual). Saldanha (2011, p. 31) define estilo como

... “um modo de traduzir” que: [1] percebe-se que pode ser reconhecido em várias traduções de um mesmo tradutor; [2] distingue o trabalho de um tradutor em relação ao trabalho de outro tradutor; [3] constitui um padrão coerente de escolhas; [4] é “motivado” no sentido de que tem uma função ou funções que se pode(m) discernir; e [5] não pode ser explicado puramente com referência ao estilo do autor ou do texto-fonte ou como resultado de restrições linguísticas.

Nota-se em Saldanha (2011) a preocupação com uma metodologia que resolva a questão ampla levantada em Baker (2000): como distinguir um estilo próprio do tradutor, consistente em várias de suas obras, independente das restrições da língua da tradução e do texto-fonte, do tipo textual e do estilo do autor?

Na subseção a seguir, são revisados brevemente os estudos do estilo do TT realizados ou em andamento no âmbito do ESTRA. Tais trabalhos, embora focalizados no estilo do TT, têm buscado conciliar as duas vertentes de pesquisa em estilo, conforme Saldanha (2011), à medida que mais textos de tradutores literários distintos passam a fazer parte do *corpus*.

1.3 Estudo do estilo do texto traduzido no âmbito do ESTRA

Barcellos (2011) e Freitas (2010) utilizam uma das possíveis composições do ESTRA, qual seja: TTs de tradutores diferentes de um mesmo TF (no caso, traduções de HOD). Barcellos (2011) analisa a apresentação do discurso (AD) – um dos traços que constituem o ponto de vista narrativo – em duas traduções publicadas em 2008, a de Sérgio Flaksman e a de José Roberto O’Shea. Esse trabalho, juntamente com Novodvorski (2013), introduz a etiquetagem das categorias de AD (*cf.* SEMINO; SHORT, 2004) e tem como principal achado a associação de dois dos traços típicos da tradução (explicitação e implicação) a mudanças no ponto de vista narrativo do texto-fonte e a prováveis estilos dos tradutores. Freitas (2010) analisa duas outras traduções de HOD, publicadas em 1984, a de Santarrita e a de Junqueira, concentrando-se na apresentação do pensamento (AP) e relacionando a explicitação e a normalização de modalidades nos TTs a possíveis mudanças no ponto de vista narrativo do TF.

Lima (2011) utiliza o conto “The dead”, da coletânea *Dubliners* (D), de James Joyce, e duas traduções, de Hamilton Trevisan e de José Roberto O’Shea, como *corpus* de estudo. Concentrando-se na apresentação da fala (AF) e especificamente no uso dos verbos *dizer/say*, o trabalho relaciona a explicitação e a normalização a mudanças do estilo do TT, associando Trevisan a um uso maior de padrões de explicitação e de normalização que O’Shea.

Novodvorski (2013) usa traduções que Sérgio Molina fez de três textos de Ernesto Sábato. Além de analisar o *corpus* a partir da etiquetagem da AD e verificar mudanças nos padrões de ocorrência das várias modalidades de AP, o autor investiga mudanças nos dêiticos espaço-temporais e pessoais e nos recursos tipográficos (*e.g.*, itálicos e pontuação) que mostram a intervenção do tradutor como mediador e mudanças no ponto de vista narrativo que constroem significados diferentes nos TTs.

Magalhães e Novodvorski (2012) e Novodvorski (2013) introduzem a investigação de palavras-chave como ponteiros para o

significado no texto e examinam padrões fraseológicos em torno dessas palavras nos TTs em comparação com aqueles do TF. Novodvorski (2013) usa, outrossim, a triangulação de dados de *corpus* (apresentação do discurso, palavras-chave e outros) com dados de paratextos e resenhas sobre o trabalho do tradutor para buscar uma explicação das mudanças significativas do estilo dos TTs.

Santos (2011) usa três contos traduzidos por O'Shea da coletânea *A good man is hard to find*, de Flannery O'Connor, a fim de verificar padrões distintos de uso da AD, com enfoque na AF e nos verbos *dizer/say*. Observa que O'Shea parece ter o propósito de aproximar os leitores do estilo do TF, mas usando, também, explicitações.

Magalhães, Castro e Montenegro (2013) analisam padrões de colocações em torno de palavras que indicam contrastes e ambiguidade na construção de significados em dois TTs de HOD, o de Fabio Cyrino em PB e o de Bernardo de Brito e Cunha em PE. As autoras observam os padrões de ocorrências e o cotexto das palavras “escuridão” e “trevas” nas traduções e apontam que parece haver mitigação, em alguns deles, ou reforço, em outros, do tema da incerteza, com prováveis funções diferentes desses padrões nas duas traduções. Finalmente, Magalhães e Blauth (no prelo) investigam o uso de itálicos, palavras estrangeiras e itens culturais específicos em seis TTs de HOD, dois portugueses e quatro brasileiros. São observados indícios de estilo de, pelo menos, dois dos tradutores brasileiros, por meio do uso singular dos traços estudados.

As contribuições desses trabalhos e de outros em andamento, usando o ESTRÁ, podem ser assim resumidas: (i) introdução da etiquetagem para análise do ponto de vista narrativo; (ii) introdução da pesquisa de palavras-chave como ponteiros para identificação dos significados do texto; (iii) investigação nos *corpora* paralelos dos resultados da análise de traços dos TTs em trabalhos com *corpora* comparáveis (sugerido por KENNY, 2005); e (iv) a triangulação de dados da análise textual do *corpus* com dados de análise contextual, por meio dos metatextos, abordados na próxima seção.

2 O ESTRA: Procedimentos de compilação, preparação e análise do *corpus*

2.1 Procedimentos de compilação

O *corpus* é compilado, de modo geral, por meio da digitalização dos textos em formato .pdf. Os arquivos resultantes dessa digitalização são, então, tratados com o programa ABBYY Fine Reader 10.0, cuja função de *Optical Character Recognition* (OCR) permite a transformação de arquivos de imagem em .pdf para outros formatos, como .doc ou .rtf. Os arquivos em novos formatos de texto são, em seguida, revisados sentença por sentença em toda a sua extensão para a correção de possíveis erros do OCR e preparação para análises (semi)automáticas usando o WordSmith Tools 6.0. O Quadro 1, a seguir, apresenta a estrutura do ESTRA até o momento.

Quadro 1: Estrutura atual do ESTRA

Nome do Arquivo	Língua	Editora	Data	Itens
ADF_Sábato	Esp.	Seix Barral	1998	31379
ADF_Molina	PB	Cia das Letras	2008	29815
AHF_Twain	Ing.	Penguin Books	1884	112180
AHF_Ferreira	PB	Casa Editora Internacional	1957	101246
AHF_Flaksman	PB	Ática	1996	115024
D_Joyce	Ing.	Grant Richards	1914	67947
D_Trevisan	PB	Civilização Brasileira	1964	59173
D_O'Sheal	PB	Siciliano	1993	69029
ET_Sábato	Esp.	Sudamericana Planeta	1948	31741
ET_Molina	PB	Cia das Letras	2008	30895
GC_Britto	PB	Cia das Letras	2011	83142
GC_Roth	Ing.	Vintage Internacional	1993	79530
GM_O'Connor	Ing.	Harvest/HBJ Book	1955	6538
GM_O'Shea	PB	Siciliano	1991	6.550

GM_Froes	PB	Cosac Naify	2008	6.431
GM_Correia	PE	Cavalo de Ferro	2008	6.650
HOD_Conrad	Ing.	Penguin Books	1994	38787
HOD_Rodrigues	PE	Dorn Quixote	1983/2009	38083
HOD_Marcos	PE	Pub. Europa-América	1999	37576
HOD_Fernandes	PE	Editorial Estampa	2006	37238
HOD_Cunha	PE	Nova Vega	2008	38133
HOD_Junqueira	PB	Itatiaia Ltda.	1984	41335
HOD_Santamita1	PB	Brasiliense	1984	36714
HOD_Trevisan	PB	Global Editora	1984	36546
HOD_Santamita2	PB	Dt Phoenix	1996	36639
HOD_Poliffr	PB	L & PM	1998	38483
HOD_Freitas	PB	Nova Alexandria	2001	38083
HOD_Paciomak	PB	Innumerus	2002	36791
HOD_Meira	PB	Martin Claret	2006	39011
HOD_Cyrino	PB	Hedra	2008	36454
HOD_Flakuzan	PB	Cia das Letras	2008	40052
HOD_O'Shea	PB	Landmark	2011	40961
IM_Britto	PB	Cia das Letras	2001	126076
IM_Lahiri	Ing.	Houghton Mifflin	1999	116016
LL_Britto	PB	Cia das Letras	2003	64417
LL_Upgrade	Ing.	Ballantine Books	2001	63074
LR_Sabato	Esp.	Seix Barral	2000	20474
LR_Molina	PB	Cia das Letras	2008	19599
TFA_Achebe	Ing.	Anchor Books	1994	41898
TFA_Silva	PB	Editora Ática SA	1983	54079
TFA_Rego	PE	Mercado de Letras	2007	50785
TOTAL				200434

O Quadro 1 apresenta a estrutura do ESTRÁ, no momento da escrita deste artigo com 41 textos já compilados e 2.004.314 itens e previsão de compilação dos demais textos até o início de 2015. Os nomes dos arquivos são constituídos: (i) pelas iniciais das palavras dos títulos dos textos (se os títulos são muito grandes, conservam-

-se apenas as iniciais de palavras lexicais); (ii) por um traço baixo; e (iii) pelo nome do tradutor ou autor.

O ESTRÁ permite a composição de diferentes tipos de *subcorpora* para estudo. Uma delas é a composição de *subcorpora* de vários TTs de um mesmo TF. Outra composição de *corpora* consiste em traduções diferentes de TFs de diferentes autores feitas por um mesmo tradutor. Em ambos os tipos de composições, investigam-se as hipóteses de Munday (2008): (i) a voz autoral do TF é fragmentada pelas várias vozes dos tradutores; e (ii) há uma padronização da diversidade de vozes autorais diferentes por razão de preferências idioletais do tradutor.

Outro tipo de composição compreende traduções de um mesmo tradutor da obra mais ampla de um autor. Esse terceiro tipo de composição é ideal para a metodologia de estilística tradutória proposta por Malmkjær (2003, 2004), que busca identificar a motivação do tradutor para criar significados diferentes nas traduções da obra mais ampla de um autor, que pode incluir tipos textuais distintos.

Atualmente, estão sendo adicionados aos textos do ESTRÁ seus paratextos (*i.e.*, orelhas, prefácios, introduções, notas do tradutor, glossários, posfácios e textos da quarta capa) – segundo parâmetros adaptados do Text Encoding Initiative (TEI), disponível em <http://www.tei-c.org/index.xml> – como objeto de estudo na busca por explicações das escolhas motivadas nos TTs. A seguir, apresentam-se os principais procedimentos de preparação do *corpus*.

2.2 Procedimentos de preparação do *corpus* para análise (semi)automática

Finda a revisão de cada texto, são adicionadas ao arquivo.doc que contém etiquetas entre parênteses angulares para marcar recursos tipográficos como itálicos, negritos – recursos esses que, juntamente com a pontuação, também podem ajudar a identificar significados e o estilo do texto e revelar comportamentos distintos dos tradutores com relação ao seu uso. A Figura 1 mostra linhas de

concordância referentes às etiquetas de itálico encontradas em excertos intercalados do TF em espanhol e suas respectivas traduções para o PB. As linhas como as apresentadas na figura são obtidas depois de converter os arquivos com extensão .doc em .txt para uso no *software* de análise lexical utilizado na pesquisa.

```
1 pronto ./s/ 15 Eso es lo que yo llamo una <itálico> buena acción <itálico> ./s/ 15 Isso é
2 ./s/ 15 Isso é o que eu chamo uma <itálico> boa ação <itálico> ./s/ 16 Piensa
3 y entonces lo obligaron a comerse una rana , <itálico> pero viva <itálico> ./s/ 19 Eu troco
4 fone e fui obrigado a comer uma rãzinha , <itálico> só que viva <itálico> ./s/ 20 No es de
5 Einstein o gente por el estilo ; respuesta : <itálico> es fácil ser modesto cuando se es
6 de Einstein ou de gente da leia ; resposta : <itálico> é fácil ser modesto quando se é
7 es célebre <itálico> ; quiero decir , <itálico> parecer modesto <itálico> ./s/ 29
8 quando se é célebre <itálico> ; quer dizer , <itálico> parecer modesto <itálico> . ./s/ 30
9 una persona que podría entenderse . <itálico> Pero fue , precisamente , la persona
10 una pessoa que poderia se entender . <itálico> Mas foi, justamente, a pessoa que
```

Figura 1: Linhas de concordância com etiqueta de itálico obtidas na análise do ESTRA

Adicionalmente, são feitos dois tipos de alinhamento dos TFs e seus TTs. O primeiro, mostrado na Figura 1, visa a que os textos fiquem intercalados, usando a sentença como unidade de alinhamento, conforme sugestão de Nélia Scott em comunicação pessoal a Jeremy Munday (*cf.* MUNDAY, 1998, p. 547). Neste caso, usa-se a ferramenta *Viewer and Aligner* do utilitário do WordSmith Tools e salvam-se os arquivos para serem lidos apenas pelo programa, com a extensão .vwr. A ideia é facilitar o processo de identificação de mudanças *in loco*. Ao se fazer a busca de uma palavra ou etiqueta no TT ou no TF, é possível, quando se amplia o texto, ver imediatamente se há mudanças na tradução que podem ser indicativas de estilo nos textos que foram intercalados.

O segundo tipo alinhamento, mostrado no Quadro 2, é feito em arquivo de extensão .docx, com os TTs e TF lado a lado. Apresenta como unidade a sentença ou o parágrafo, dependendo do foco de análise.

Quadro 2: Tela com exemplo de alinhamento lado a lado

O'CONNOR	O'SHEA	FROES
THE PEACOCK was following Mrs Shortley up the road to the hill where she meant to stand.	O pavão seguia a sra. Shortley estrada acima em direção à colina onde ela pretendia se posicionar.	Estrada acima, o pavão seguia Mrs Shortley para o morrinho onde ela pretendia ficar.
Moving one behind the other, they looked like a complete procession.	Seguindo um atrás do outro, pareciam uma perfeita procissão.	E os dois, andando um atrás do outro, já pareciam uma procissão por inteiro.
Her arms were folded and as she mounted the prominence, she might have been the giant wife of the countryside, come out at some sign of danger to see what the trouble was.	Os braços dela estavam cruzados e, subindo a encosta, parecia a esposa gigante do campo, que, ao sinal de perigo, viera investigar o problema.	Ao galgar a elevação, aquela mulher, que ia de braços cruzados, bem poderia ter sido a colossal esposa do campo, saída a algum sinal de perigo para ver qual era o problema.

Outros procedimentos, descritos na subseção a seguir, envolvem, simultaneamente, análise e preparação.

2.3 Procedimentos de análise

Os procedimentos de análise variam de acordo com o objeto de estudo dentro do amplo espectro do tema “estilo do texto”. Um deles é, após levantar a lista de palavras, verificar se aquelas que ocorrem apenas uma vez podem ser consideradas *hapax legomena* (i.e., indícios de criatividade no texto). Outro é usar a lista de palavras-chave e examiná-las quanto à sua chavicidade nos TTs em relação aos TFs. Linhas de concordâncias podem ser feitas usando palavras com alto nível de chavicidade como nódulos para verificar seus colocados, agrupamentos lexicais em torno delas e seus *plots*,

tentando mapear prováveis regularidades que indiquem relevância literária ou destaque também nos TTs. Em outra possibilidade ainda de análise, linhas de concordância são feitas com as etiquetas de itálico ou palavras estrangeiras como nódulos de busca para examinar o uso desses recursos nos TTs.

Alguns procedimentos de análise estão estreitamente relacionado a outros, de preparação do texto, e vice versa. Por exemplo, classificam-se as ocorrências de AD, ou seja, modalidades de AF (cf. Quadro 3), de AP e de apresentação da escrita (AE) – de acordo com o que propõem Semino e Short (2004) –, etiquetando-as e inserindo-as entre parênteses angulares no texto que está sendo preparado.

Quadro 3: Categorias de AF

Apresentação da Fala	
<RFN>	Relato de Fala pelo Narrador
<NV>	Narração de Voz
<RAFN>	Relato de Ato de Fala pelo Narrador
<FI>	Fala Indireta
<FIL>	Fala Indireta Livre
<FD>	Fala Direta
<FDL>	Fala Direta Livre

A AD também pode ser examinada com a busca de sinais de pontuação como as aspas para as modalidades diretas e do “que” e do “that” para algumas das modalidades indiretas. Trata-se de elementos que podem ser esclarecedores sobre o ponto de vista narrativo. O tradutor, no TT, faz o papel do narrador, mediador entre os personagens e o leitor, podendo fazer escolhas distintas, quer da modalidade de AD, quer do verbo usado na oração introdutória de AD. Com isso, o tradutor, ao mesmo tempo em que interpreta o discurso da personagem, pode vir a apresentar um ponto de vista dessa personagem que é mais ou menos explícito em relação àquele do TF.⁴

Considerações finais

Este artigo teve como objetivo mostrar os desdobramentos da linha de pesquisa de estilo da tradução/tradutor orientada para *corpus* no contexto nacional. Para tanto, fez um relato de trabalhos iniciais desenvolvidos usando a metodologia de *corpus* e uma abordagem linguística da tradução; em seguida, relatou trabalhos em que a questão do estilo era latente; finalmente, chegou a uma fase em que é adotada, principalmente, a abordagem do estilo da tradução, mas tentando integrar estudos também do estilo de tradutores – abordagem essa que resultou na criação do ESTRA. Ao descrever os procedimentos de compilação, preparação e análise do ESTRA, este artigo procurou mostrar as contribuições originais, seja na compilação do *corpus*, na preparação ou nas metodologias usadas para análise do estilo. Não foi possível, dado o escopo do artigo, apresentar análises de aspectos do estilo usando exemplos dos diferentes *corpora*, como feito em trabalhos anteriores (cf. MAGALHÃES; NOVODVORSKI, 2012; MAGALHÃES; CASTRO; MONTENEGRO, 2013 e MAGALHÃES; BLAUTH, no prelo). Por ora, vale salientar que constituem perspectivas futuras: (i) finalizar a compilação do ESTRA; (ii) finda essa compilação, usar os metatextos detalhadamente para realizar a etapa da explicação, por meio de fatores extralinguísticos, seja da motivação do tradutor para as escolhas deliberadas, seja da agenda ideológica do tradutor expressa em escolhas mais ou menos conscientes, seja de sua voz ou presença discursiva na tradução, incluindo seu papel no processo de escolha do que traduzir e de edição da tradução. Os metatextos do ESTRA podem incluir, ainda, as capas dos TTs, cujas imagens são escolhas potencialmente significativas quanto ao propósito e ao público-alvo da tradução. Os metatextos permitirão a consolidação de uma metodologia de triangulação dos dados linguísticos com dados extralinguísticos, o que contribuirá para a confirmação ou não das hipóteses levantadas por Munday (2008) e outras que possam ser elaboradas no curso do trabalho sobre do estilo do texto/tradutor.

Notas

1. Este artigo é resultado de pesquisa financiada pela FAPEMIG (PPM-00020-10, CNPq (PQ 302178/2010-4) e CAPES (PACCCSSII 025/2013) e realizada no âmbito do grupo de pesquisa ESTRALI, do CNPq.
2. O escopo deste artigo não permite uma revisão de trabalhos afiliados à linha em tela em outros centros de pesquisa brasileiros, mesmo reconhecendo seu significativo número e relevância.
3. Do original *priming* (cf. HOEY, 2005).
4. Ver, por exemplo, em Lima (2011), as escolhas feitas por Hamilton Trevisan para traduzir os verbos das orações introdutórias das falas dos personagens – escolhas essas distintas das que faz José Roberto O’Shea.

Referências

AGUIAR, S.M. *As vozes de Chico Buarque em inglês: tradução e linguística de corpus*. 2010. 201f. Tese (Estudos Linguísticos e Literários em Inglês) – FFLCH/USP, São Paulo, 2010.

ASSIS, R.C. *A transitividade na representação de Sethe no corpus paralelo Beloved/Amada*.

2004. 122f. Dissertação (Linguística Aplicada) FALE/UFMG, Belo Horizonte, 2004.

_____. *A representação de europeus e africanos como atores sociais em Heart of Darkness (O coração das trevas) e em suas traduções para o português: uma*

abordagem textual da tradução. 2009. 267f. Tese (Linguística Aplicada) – FALE/UFMG, Belo Horizonte, 2009.

BAKER, M. Corpus linguistics and translation studies: implications and applications. In: BAKER, B.; FRANCIS, G.; TOGNINI-BONELLI, E. (Ed.). *Text and technology*. Amsterdã: John Benjamins, 1993. p. 233-250.

_____. *Corpora in translation studies: an overview and some suggestions for future research*. *Target*, v. 7, n. 2, p. 223-243, 1995.

_____. Towards a methodology for investigating the style of a literary translator. *Target*, v. 12, n. 2, p. 241-266, 2000.

BARCELLOS, C.P. *O estilo de tradutores: apresentação do discurso no corpus paralelo Heart of darkness/(No) Coração das trevas*. 2011. 154f. Dissertação (Linguística Aplicada) – FALE/UFMG, Belo Horizonte, 2011.

BUENO, L.T. *Transitividade, coesão e criatividade lexical no corpus paralelo Macunaíma, de Andrade e Macunaíma, de Goodland*. 2005. 176f. Dissertação (Linguística Aplicada) – FALE/UFMG, Belo Horizonte, 2005.

_____. *Chapeuzinho Vermelho e Caperucita Roja: uma investigação de reescritas com base na representação (visual) de atores sociais e na representação da ação social*. 2011. 296f. Tese (Linguística Aplicada) – FALE/UFMG, Belo Horizonte, 2011.

CAMARGO, D.C. Uma análise de semelhanças e diferenças na tradução de textos técnicos, jornalísticos e literários. *D.E.L.T.A.*, v. 20, n. 1, p. 1-25, 2004

FREITAS, G.P. *Estilo em tradução: um estudo do ponto de vista narrativo e atos de fala e pensamento no corpus paralelo Heart of darkness/Coração das trevas*. 2010. 51f. TCC (Graduação em Letras) – FALE/UFMG. Belo Horizonte, 2010.

HALLIDAY, M.A.K. Linguistic function and literary style: an inquiry into the language of William Golding's *The Inheritors*. In: CHATMAN, S. (Ed.) *Literary style: a symposium*. Londres: OUP, 1971. p. 330-368.

HALLIDAY, M.A.K. *An introduction to functional grammar*. 2. ed. Londres: Edward Arnold, 1994.

HATIM, B.; MASON, I. *Discourse and the translator*. Londres: Longman, 1992.

HOEY, M. *Lexical priming: a new theory of word and language*. Londres: Routledge, 2005.

KENNY, D. *Lexis and creativity in translation: a corpus-based study*. Manchester: St. Jerome Publishing, 2001.

_____. Parallel *corpora* and translation studies: old questions, new perspectives? Reporting *that* in Gepcolt: a case study. In: BARNBOOK, G.; DANIELSSON, P.; MAHLBERT, M. (Ed.) *Meaningful texts: the extraction of semantic information from monolingual and multilingual corpus*. Londres: Continuum, 2005. p. 154-165.

LEECH, G.; SHORT, M. *Style in fiction: a linguistic introduction to English fictional prose*. Harlow: Pearson Education, [1981]2007.

LIMA, C. M. *Estilo em tradução literária: um estudo da presença do tradutor nos relatos de atos de fala do corpus paralelo *The dead/Os mortos**. 2011. 41f. TCC (Graduação em Letras) – FALE/UFMG. Belo Horizonte, 2011.

MAGALHÃES, C.M. Da coesão como recurso de continuidade do discurso. In: ALVES, F.; MAGALHÃES, C.M.; PAGANO, A. (Org.). *Competência em tradução: cognição e discurso*. Belo Horizonte: UFMG, 2005. p. 209-245.

_____; ASSIS, R.C. Representação de atores sociais em corpus paralelo: *Heart of darkness* e suas traduções para o português. In: COHEN, M.A.; LARA, G.M.P. (Org.). *Linguística, tradução, discurso*. Belo Horizonte: UFMG, 2009. v. 1, p. 201-220.

_____; BLAUTH, T.P. Estilo do tradutor: um estudo do uso de itálicos, palavras estrangeiras e itens culturais específicos por seis tradutores do português de *Heart of Darkness*. In: VIANA, V.; TAGNIN, S.E.O. *Corpora na tradução. (no prelo)*

_____; CASTRO, M.C.; MONTENEGRO, M.S. Estilística tradutória: um estudo de cópua paralelo de uma tradução brasileira e uma tradução portuguesa de *Heart of darkness*. *TradTerm*, v. 21, p. 11-29, jul. 2013.

_____; NOVODVORSKI, A. A chavicidade na análise de estilo em tradução: um estudo baseado em *corpora* paralelos espanhol/português. In: DUTRA, D. P.; MELLO, H. (Org.). *Anais do X Encontro de Linguística de Corpus: aspectos metodológicos dos estudos de corpora*. Belo Horizonte: FALE/UFMG, 2012. p. 294-313.

MALMKJÆR, K. Love thy neighbour: will parallel *corpora* endear linguists to translators? *Meta*, v. 43, n. 4, p. 534-541, 1998.

_____. What happened to God and the angels? An exercise in translational stylistics. *Target*, v. 15, n. 1, p. 37-58, 2003.

_____. Translational stylistics: Dulcken's translations of Hans Christian Andersen. *Language and Literature*, v. 13, n. 1, p. 13-24, 2004.

_____. Linguistic approaches to translation. In: MALMKJÆR, K.; WINDLE, K. (Ed.). *The Oxford handbook of translation studies*. Nova York: OUP, 2011. p. 57-70.

_____; CARTER, R. Stylistics. In: MALMKJÆR, K. (Ed.) *The linguistics encyclopedia*. 2. ed. Londres: Routledge, 2002. p. 510-520.

MAURI, C. *Um estudo da tradução italiana de Laços de Família, de Clarice Lispector, a partir da abordagem em corpora: a construção da introspecção feminina através dos verbos de elocução*. 2003. 101f. Dissertação (Linguística Aplicada) – FALE/UFMG, Belo Horizonte, 2003.

_____. *Uma análise do ponto de vista em A hora da estrela e Laços de família, de Clarice Lispector, e nas traduções italianas L'ora della stella e Legami Familiari*. 2009. 258f. Tese (Linguística Aplicada) – FALE/UFMG. Belo Horizonte, 2009.

MUNDAY, J. A computer-assisted approach to the analysis of translation shifts. *Meta*, v. 43, n. 4, p. 542-556, 1998.

_____. Translation and ideology: a textual approach. *The Translator*, v. 13, n. 2, p. 195-217, 2007.

_____. *Style and ideology in translation: Latin American writing in English*. Londres: Routledge, 2008.

NIDA, E.A. *Toward a science of translating*. Leiden: E. J. Brill, 1964.

NOVODVORSKI, A. *Estilo das traduções de Sergio Molina de obras de Ernesto Sabato: um estudo de corpora paralelos espanhol/português*. 2013. 259f. (Doutorado em Linguística Aplicada) – FALE/UFMG, Belo Horizonte, 2013.

SALDANHA, G. Translator style: methodological considerations. *The Translator*, v. 17, n. 1, p. 25-50, 2011.

SANTOS, P. *Estilo em tradução: um estudo dos atos de fala no corpus paralelo A good man is hard to find/É difícil encontrar um homem bom*. 2011. 45f. TCC (Graduação em Letras) – FALE/UFMG. Belo Horizonte, 2011.

SCOTT, M. *Normalisation and readers' expectations: a study of literary translation with reference to Clarice Lispector's A hora da estrela*. 1998. 331f. Tese (Tradução) – Faculty of Humanities and Social Sciences, University of Liverpool, Liverpool, 1998.

SEMINO, E.; SHORT, M. *Corpus stylistics: speech, writing and thought presentation in a corpus of English writing*. Londres: Routledge, 2004.

SIMPSON, P. *Language, ideology and point of view*. Londres: Routledge, 2000.

VINAY, J.-P.; DARBELNET, J. *Comparative stylistics of French and English: a methodology for translation*. Amsterdã: John Benjamins, [1958]1995.

ZANETTIN, F. *Translation-driven corpora*. Manchester: St. Jerome, 2012.

Recebido: 19-02-14

Aceito: 25-04-14